

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/581,158
Source: FGBP
Date Processed by STIC: 6/9/06

ENTERED



IFWP

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006
TIME: 10:35:46

Input Set : A:\sequence listing.DOC
Output Set: N:\CRF4\06092006\J581158.raw

3 <110> APPLICANT: Yoshinori Watanabe
 5 <120> TITLE OF INVENTION: Novel centromeric protein SHUGOSHIN
 7 <130> FILE REFERENCE: 4439-4043
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/581,158
 C--> 9 <141> CURRENT FILING DATE: 2006-05-31
 9 <150> PRIOR APPLICATION NUMBER: JP2003-401943
 10 <151> PRIOR FILING DATE: 2003-12-01
 12 <150> PRIOR APPLICATION NUMBER: JP2004-279450
 13 <151> PRIOR FILING DATE: 2004-09-27
 15 <160> NUMBER OF SEQ ID NOS: 45
 17 <170> SOFTWARE: PatentIn version 3.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 960
 21 <212> TYPE: DNA
 22 <213> ORGANISM: yeast
 24 <400> SEQUENCE: 1
 25 atgaactttc aatttataaa ttcaaatata aacaatgaag ataaattgcc gatggagtcg 60
 27 ttgaaaaaga aattttaaa acaaaatcgta gaaattataa aaataaaatac tcagctttct 120
 29 ataaaaatta gagaatctga aaacgaaaatt caagatttga tacaagaaaa tttcactttg 180
 31 aaaagttatt tggtaaact tgaagctcgta ttgcgaatc aatctcaaacc tgaggacttg 240
 33 taaaaaaact tcttcctgta gatacaaacc attcacaaaa agatttcaca agtgc当地 300
 35 ttactgaaga ttatagagaa aaagtgttca tcagatttcc tcgaagcgaat tgaaaaaaat 360
 37 caatttacaa cctgtgaaaa taaagattcg aaagaagatt atcagatttt gcataataaa 420
 39 cgcttggagt atgtatcatt taatgtatgaa cttaaaagtc tcgaaacagg gcaaccattg 480
 41 tattgttttc aagatttcca aaaaaaaatgc catggccctc cggctctatc tgaaaaacct 540
 43 ggaaaatgta tattaaaaaga taaaaccaat gcccacgtaa acaaaaatacc acaagatgag 600
 45 gtgaattact cattgcgcga aaaaaatatc accatcttt caaaggaaatt aaaagaaaaac 660
 47 gaatttgaat ccatcaacga gggcggaaact gaagaagaaa aggctaaaac atcaaattgtt 720
 49 tgtgtttgtta ttccctgttaa aagtgtgttca cagataactg accttaaagg acaagcaacc 780
 51 ggagacagct ccccatgtga tttgaagaa tctcaaccaa ggattatgg acgtgaaaaaa 840
 53 ctaagacgat cagtcaaaatg gataaaactat gcaataacca gtttgcgaac taaaactacga 900
 55 cgagactttg acttaccatc tgatagaaaa cgcaaacgac atcccagagg caaagcataaa 960
 58 <210> SEQ ID NO: 2
 59 <211> LENGTH: 319
 60 <212> TYPE: PRT
 61 <213> ORGANISM: yeast
 63 <400> SEQUENCE: 2
 65 Met Asn Phe Gln Phe Ile Asn Ser Asn Ile Asn Asn Glu Asp Lys Leu
 66 1 5 10 15
 69 Pro Met Glu Ser Leu Lys Lys Lys Phe Leu Lys Gln Asn Arg Glu Ile
 70 20 25 30
 73 Ile Lys Ile Asn Thr Gln Leu Ser Ile Lys Ile Arg Glu Ser Glu Asn
 74 35 40 45

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006
TIME: 10:35:46

Input Set : A:\sequence listing.DOC
Output Set: N:\CRF4\06092006\J581158.raw

77 Glu Ile Gln Asp Leu Ile Gln Glu Asn Phe Thr Leu Lys Ser Tyr Leu
 78 50 55 60
 81 Val Lys Leu Glu Ala Arg Phe Arg Asn Gln Ser Gln Thr Glu Asp Leu
 82 65 70 75 80
 85 Leu Lys Asn Phe Pro Glu Ile Gln Thr Ile His Lys Lys Ile Ser
 86 85 90 95
 89 Gln Val Gln Ser Leu Leu Lys Ile Ile Glu Lys Lys Cys Ser Ser Asp
 90 100 105 110
 93 Phe Leu Glu Ala Asn Val Lys Ser Gln Phe Thr Thr Cys Glu Asn Lys
 94 115 120 125
 97 Asp Ser Lys Glu Asp Tyr Gln Ile Leu His Asn Lys Arg Leu Glu Tyr
 98 130 135 140
 101 Val Ser Phe Asn Asp Glu Leu Lys Ser Leu Glu Thr Gly Gln Pro Leu
 102 145 150 155 160
 105 Tyr Cys Phe Gln Asp Phe Gln Lys Lys Val His Gly Pro Pro Ala Leu
 106 165 170 175
 109 Ser Glu Lys Pro Gly Lys Cys Ile Leu Lys Asp Lys Thr Asn Ala His
 110 180 185 190
 113 Val Asn Lys Ile Pro Gln Asp Glu Val Asn Tyr Ser Leu Pro Gln Lys
 114 195 200 205
 117 Asn Ile Thr Ile Phe Ser Lys Glu Leu Lys Glu Asn Glu Phe Glu Ser
 118 210 215 220
 121 Ile Asn Glu Gly Glu Thr Glu Glu Lys Ala Lys Thr Ser Asn Val
 122 225 230 235 240
 125 Cys Val Cys Ile Pro Cys Lys Ser Ala Glu Gln Ile Thr Asp Leu Lys
 126 245 250 255
 129 Gly Gln Ala Thr Gly Asp Ser Ser Pro Cys Asp Phe Glu Glu Ser Gln
 130 260 265 270
 133 Pro Arg Ile Asn Gly Arg Glu Lys Leu Arg Arg Ser Val Lys Val Ile
 134 275 280 285
 137 Asn Tyr Ala Ile Pro Ser Leu Arg Thr Lys Leu Arg Arg Asp Phe Asp
 138 290 295 300
 141 Leu Pro Ser Asp Arg Lys Arg His Pro Arg Gly Lys Ala
 142 305 310 315
 145 <210> SEQ ID NO: 3
 146 <211> LENGTH: 1944
 147 <212> TYPE: DNA
 148 <213> ORGANISM: yeast
 150 <400> SEQUENCE: 3
 151 atgtcgaaag catcttttc cccgaacgta gaagacttga aaaaaaaagca aattcgacag 60
 153 tataaggaaa ttatacgaat aagcaaggca caatcaatta gaattaaaga attgcagtta 120
 155 gaaaatgaac ggttgtttc ggaaaaatatc gatttgagga ctacagcgat aaacttggaa 180
 157 gagcaactcg aaaccgtgca aaacgaaaaac gaagaaaaca aaacaaagtt agctgcatta 240
 159 cttaatcgat ttcatgaaga aacagataat ttttatcaa aattaagtct ttgtcagcaa 300
 161 gaaataacaag acaccttcaa accagtggag gctaacttag cttacgatgt cgatacggat 360
 163 tctgaagacc ttgacgagga atccgtcgtg aaagataccg aagaaataat tgagcaagct 420
 165 cagcatgatg tttccttacg aaatttaagt ggaatagagg atgaaaatat aattgatgac 480
 167 ggagaaaactg ctataaatga aaaaaaaaaa agagaagcta atgtttttc cgacacgcaa 540
 169 tcagcacctc agctaaaatc cggcaagcc ctcccagctg attttgaaaa tccttacaat 600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006

TIME: 10:35:46

Input Set : A:\sequence listing.DOC

Output Set: N:\CRF4\06092006\J581158.raw

171 ctatccaatt cgaaacctgt aaataataat aatgaagata gagttgaagc ggttacttct 660
 173 gaaaataaat ctatcgattc tgctcctcag gaaaaaaatc atgaatacga aatcgtagt 720
 175 ccaaaatcat tatccaacaa aattaataat caagcagctg cacaagaag aaccgaagaa 780
 177 gataatgcaa atggagggtgc tcaagaagaa aatgagggtt cacaagaagc tcattttcat 840
 179 agcagaatac aatctgatac agtaatacaa agtacacccca ctaaacggaa atggacgtt 900
 181 gacattcaaa ataaacaaat taatctggct tctcagcta ccaatgttac cggttatgt 960
 183 tcggagaccg atagtcgccc caatcgcgca aactcttgg attctgctgt ccttcttg 1020
 185 caatctcaa ataaaagtaa cgaaaatggg catcatattt cagatcccaa tttaaatagc 1080
 187 tccatatcgt tgaagttgc gcctgaagat actgcgcata attcattaac ttcacaagag 1140
 189 aatgttggc ctcagggtac gacgacttct ctgtcaaata tgactgttc tgaatctct 1200
 191 cgtacagaca ctccaaggaa aataaacggg ttagtagact cttctgtcac taatggaaac 1260
 193 gaaaaatttt ctgtagaaat aatgaatgac tctaacaaaa ttggactgaa tcctaaatct 1320
 195 tttaccgacg aagagcgggaa aattttaaaca cttttcgaa atcctcccat gagactgtca 1380
 197 agtgaacctc catcttcaa tggattttca atagcccatc ccaataattc tccggtacgt 1440
 199 ccgccatcgc tacaaggaaat attgaatgct gaagatcgac cttacgaaat tgagccgtca 1500
 201 cgtagctcct ttgctaccaa cgatacgggc tcctataata atttggact tctgtcatct 1560
 203 gtaacgaatt taaaatcccc taatgagaac gatcgtgtga cggaaaactca gtcgcgaaga 1620
 205 gaaacaaaag taaaaggcg aagaaaaagct cggattcaag aaacttctga agaaagtaca 1680
 207 gtagtcaatg agccaaatga aaaacctgat ggaaggagcc gaagggaaacg gaaaaagggtt 1740
 209 aattacgctt tgcctggatt aaggacgaaa ttaagacgga atttcgattt accttcagat 1800
 211 catgtaaaag ctaaaaaaac gagacgtgct cctaagaact ctgagaatga ttcagctacc 1860
 213 aaaacagaaaa ccgcaaacat tacttctgaa gcacccacta cttcagaagt aacccttgaa 1920
 215 aactccgaaa cccttaattt gtaa 1944
 218 <210> SEQ ID NO: 4
 219 <211> LENGTH: 647
 220 <212> TYPE: PRT
 221 <213> ORGANISM: yeast
 223 <400> SEQUENCE: 4
 225 Met Ser Lys Ala Ser Leu Ser Pro Asn Val Glu Asp Leu Lys Lys Lys
 226 1 5 10 15
 229 Gln Ile Arg Gln Tyr Lys Glu Ile Ile Arg Ile Ser Lys Ala Gln Ser
 230 20 25 30
 233 Ile Arg Ile Lys Glu Leu Gln Leu Glu Asn Glu Arg Leu Leu Ser Glu
 234 35 40 45
 237 Asn Ile Asp Leu Arg Thr Thr Ala Ile Asn Leu Glu Glu Gln Leu Glu
 238 50 55 60
 241 Thr Val Gln Asn Glu Asn Glu Glu Asn Lys Thr Lys Leu Ala Ala Leu
 242 65 70 75 80
 245 Leu Asn Arg Phe His Glu Glu Thr Asp Asn Phe Leu Ser Lys Leu Ser
 246 85 90 95
 249 Leu Cys Gln Gln Glu Ile Gln Asp Thr Phe Lys Pro Val Glu Ala Asn
 250 100 105 110
 253 Leu Ala Tyr Asp Val Asp Thr Asp Ser Glu Asp Leu Asp Glu Glu Ser
 254 115 120 125
 257 Val Val Lys Asp Thr Glu Glu Ile Ile Glu Gln Ala Gln His Asp Val
 258 130 135 140
 261 Ser Leu Arg Asn Leu Ser Gly Ile Glu Asp Glu Asn Ile Ile Asp Asp
 262 145 150 155 160
 265 Gly Glu Thr Ala Ile Asn Glu Gln Lys Lys Arg Glu Ala Asn Val Phe

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006
TIME: 10:35:46

Input Set : A:\sequence listing.DOC
Output Set: N:\CRF4\06092006\J581158.raw

266	165	170	175	
269	Ser Asp Thr Gln Ser Ala Pro Gln Leu Lys Ser Gly Lys Ala Leu Pro			
270	180	185	190	
273	Ala Asp Phe Glu Asn Pro Tyr Asn Leu Ser Asn Ser Lys Pro Val Asn			
274	195	200	205	
277	Asn Asn Asn Glu Asp Arg Val Glu Ala Val Thr Ser Glu Asn Lys Ser			
278	210	215	220	
281	Ile Asp Ser Ala Pro Gln Glu Lys Asn His Glu Tyr Glu Ile Val Ser			
282	225	230	235	240
285	Pro Lys Ser Leu Ser Asn Lys Ile Asn Asn Gln Ala Ala Ala Gln Arg			
286	245	250	255	
289	Arg Thr Glu Glu Asp Asn Ala Asn Gly Val Ala Gln Glu Glu Asn Glu			
290	260	265	270	
293	Gly Ser Gln Glu Ala His Phe His Ser Arg Ile Gln Ser Asp Thr Val			
294	275	280	285	
297	Ile Gln Ser Thr Pro Thr Lys Arg Lys Trp Asp Val Asp Ile Gln Asn			
298	290	295	300	
301	Lys Gln Ile Asn Leu Ala Ser Ala Ala Thr Asn Val Thr Gly Tyr Val			
302	305	310	315	320
305	Ser Glu Thr Asp Ser Arg Pro Asn Arg Ala Asn Ser Leu Asp Ser Ala			
306	325	330	335	
309	Val Leu Leu Val Gln Ser Ser Asn Lys Ser Asn Arg Asn Gly His His			
310	340	345	350	
313	Ile Ser Asp Pro Asn Leu Asn Ser Ser Ile Ser Leu Lys Phe Ala Pro			
314	355	360	365	
317	Glu Asp Thr Ala His Asn Ser Leu Thr Ser Gln Glu Asn Val Gly Pro			
318	370	375	380	
321	Gln Val Thr Thr Ser Leu Ser Asn Met Thr Val Ala Glu Ser Pro			
322	385	390	395	400
325	Arg Thr Asp Thr Pro Arg Glu Ile Asn Gly Leu Val Asp Ser Ser Val			
326	405	410	415	
329	Thr Asn Gly Asn Glu Lys Phe Ser Val Glu Ile Met Asn Asp Ser Asn			
330	420	425	430	
333	Lys Ile Gly Leu Asn Pro Lys Ser Phe Thr Asp Glu Glu Arg Glu Ile			
334	435	440	445	
337	Leu Thr Leu Phe Arg Asn Pro Pro Met Arg Leu Ser Ser Glu Pro Pro			
338	450	455	460	
341	Ser Ser Asn Gly Phe Ser Ile Ala His Pro Asn Asn Ser Pro Leu Arg			
342	465	470	475	480
345	Pro Pro Ser Leu Gln Gly Ile Leu Asn Ala Glu Asp Arg Pro Tyr Glu			
346	485	490	495	
349	Ile Glu Pro Ser Arg Ser Ser Phe Ala Thr Asn Asp Thr Gly Ser Tyr			
350	500	505	510	
353	Asn Asn Leu Glu Leu Leu Ser Ser Val Thr Asn Leu Lys Ser Pro Asn			
354	515	520	525	
357	Glu Asn Asp Arg Val Thr Lys Thr Gln Ser Arg Arg Glu Thr Lys Val			
358	530	535	540	
361	Lys Arg Arg Arg Lys Ala Arg Ile Gln Glu Thr Ser Glu Glu Ser Thr			
362	545	550	555	560

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006

TIME: 10:35:46

Input Set : A:\sequence listing.DOC
 Output Set: N:\CRF4\06092006\J581158.raw

```

365 Val Val Asn Glu Pro Asn Glu Lys Pro Asp Gly Arg Ser Arg Arg Glu
366           565           570           575
369 Arg Lys Lys Val Asn Tyr Ala Leu Pro Gly Leu Arg Thr Lys Leu Arg
370           580           585           590
373 Arg Asn Phe Asp Leu Pro Ser Asp His Val Lys Ala Lys Lys Thr Arg
374           595           600           605
377 Arg Ala Pro Lys Asn Ser Glu Asn Asp Ser Ala Thr Lys Thr Glu Thr
378           610           615           620
381 Ala Asn Ile Thr Ser Glu Ala Pro Thr Thr Ser Glu Val Thr Leu Glu
382 625           630           635           640
385 Asn Ser Glu Thr Leu Asn Leu
386           645
389 <210> SEQ ID NO: 5
390 <211> LENGTH: 1773
391 <212> TYPE: DNA
392 <213> ORGANISM: yeast
394 <400> SEQUENCE: 5
395 atgccgaaga gaaaaattgc tcctaacaag gaaagcagca ggcgtacgg ctcccacgat      60
397 gatttaaccc cacaataaca agaatttcaa aacctaattgg atctcgaaatc gcaaaaagtg    120
399 gaaaacatca gacagtcgta ttcgaggcaa aactccctgc tggccaagga taactccata    180
401 taaaaaatta aagttaatag cttggaaaaaa aaaataagcc agctggtaca agaaaacgtg    240
403 actctacgat ctaaaacctc tataagcgaa gctatctaca gggAACGGTT aagtaatcaa    300
405 ctacaagtca ttgaaaacgg tattattcaa agatttgacg aaattttta tatgttttag     360
407 aacgtacgta aaaacgaaaaa ttgccccagt tcgagctaa gaacaatgtt gaagagaacg    420
409 agttccaggt caagatcatg ctcattgtca tcaccacat actcaaaaaag ttacactagg   480
411 ttatcaaatc acgagaataa cctgtcgcat gaatcaagtt ttaacaagga cgatggtcca   540
413 gatcttgagc ctaaggctaa aaaaaggaag agttcttaggc ggcaatctat gtttgtatcc  600
415 acgagtttag aacctgaaga cgaaaccggt gaaaacgaac ccatgatgga aaattccctct  660
417 gtagaggtac cggcagaatc acacgagttc ggcgaagtgg agggaaacaat agatgcctta 720
419 aaccctgaag agggaaaatag cgattctgtc agtaatttta ccaattcaat tatagaatac  780
421 tccataccag aggagaatcc gacagaaccc gagcattcat cttctaaact agaaatattc  840
423 aatgacagta caaatatgct aagtacagtg cctgtcaaatc ctttgcgtt gcctttacca 900
425 gccccatccg caactttacc tactaccact agcgatgtt caacgggtca tccttcatca 960
427 agttcttcta ctaattctca tccaaagacc aaaattaagc attccatgaa gccgcctagg 1020
429 atagaactga agaaaaaggt tattgacgaa gtcatgcccgt taagtaacat gagcagcaac 1080
431 agcgaaatat catttacgag aactagaaga actcggtt aagctgtaga ttacactttg 1140
433 ctttctttaa gagccaaaat gaggaggcct tcagaaaaac ttgtggatgc tactactgtg 1200
435 attgatatac atgatctaca gttttccaag agaaatcggtt aaacttcaca taaaaggaaa 1260
437 agtttatccc aagattcaat acccgacgaa ccgcattgtt gagaagtgcgt cgctctaaag 1320
439 gattatggaa ctccaaaagg gaaaaaaaaacg gaagatggaa tacacgagga taccgctcat 1380
441 ctaatgacca cttccaaacaa caacagcaac aacaaaaacg aaaaaaaaaact aactagcaac 1440
443 aatagcccta aaaaatcgtc gcctttactt gacattacaa ataaatcggtt gaataagaaa 1500
445 aagtcaacaa gaactaaaaa attgttcaaa aatgcaatttgc tcaataattt atctgtatgg 1560
447 aattctacta cgcgaccctc caagtcgttca aaggaaacca gtaataataa caacaattac 1620
449 aacaatttcg acaataacaa ttcaaaccatt aataatgtt aataataatc tgttagcttt 1680
451 agactaaaatg aagatgattt agcagtattt gatttattt gaaatggtaa ggcagtgaaa 1740
453 catcaaccaa aaacatatcg caccaaaaaa tga                           1773
456 <210> SEQ ID NO: 6
457 <211> LENGTH: 590

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/09/2006
PATENT APPLICATION: US/10/581,158 TIME: 10:35:47

Input Set : A:\sequence listing.DOC
Output Set: N:\CRF4\06092006\J581158.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:40

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/581,158

DATE: 06/09/2006

TIME: 10:35:47

Input Set : A:\sequence listing.DOC

Output Set: N:\CRF4\06092006\J581158.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date